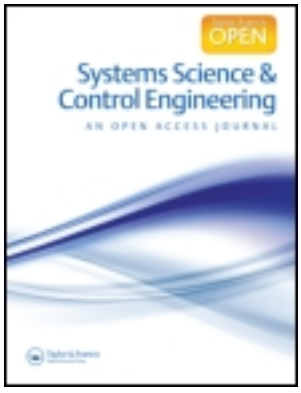
On: 19 May 2014, At: 14:21
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Systems Science & Control Engineering: An Open Access Journal

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tssc20

Online suboptimal control of linearized models

V. Costanza[a] & P.S. Rivadeneira[a]

[a] "Grupo de Sistemas No Lineales", INTEC-Facultad de Ingeniería Química (UNL-CONICET), Güemes 3450, 3000 Santa Fe, Argentina

Accepted author version posted online: 08 Apr 2014.Published online: 08 Apr 2014.

To cite this article: V. Costanza & P.S. Rivadeneira (2014) Online suboptimal control of linearized models, Systems Science & Control Engineering: An Open Access Journal, 2:1, 379-388, DOI: 10.1080/21642583.2014.913215

To link to this article: http://dx.doi.org/10.1080/21642583.2014.913215

Systems Science & Control Engineering: An Open Access Journal, 2014
Vol. 2, 379–388, http://dx.doi.org/10.1080/21642583.2014.913215

Taylor & Francis
Taylor & Francis Group

Online suboptimal control of linearized models

V. Costanza* and P.S. Rivadeneira

"Grupo de Sistemas No Lineales", INTEC-Facultad de Ingeniería Química (UNL-CONICET), Güemes 3450, 3000 Santa Fe, Argentina

(*Received 20 November 2013; final version received 5 April 2014*)

A novel approach to approximately solving the restricted-control linear quadratic regulator problem online is substantiated and applied in two case studies. The first example is a one-dimensional system whose exact solution is known. The other one refers to the temperature control of a metallic strip at the exit of a multi-stand rolling mill. The new (online-feedback) strategy employs a convenient version of the gradient method, where partial derivatives of the cost are taken with respect to the final penalization matrix coefficients and to the switching times where the control (de)saturates. The calculations are based on exact algebraic formula, which do not involve trajectory simulations, and so reducing in principle the computational effort associated with receding horizon or nonlinear programming methods.

Keywords: optimal control; restricted controls; linear quadratic regulator problem; online optimization

1. Introduction

Linearized models are frequently employed to treat nonlinear systems evolving near an equilibrium point, or when tracking a reference trajectory. These approximate models are accepted provided both deviations (of the state from the given target and of the manipulated variable from the reference control) are 'small'. Therefore, restrictions on the control values appear naturally when working with linearizations. In most cases it is implicitly assumed that the linear approximation is BIBO stable and controllable as to guarantee that if the control moves between the imposed bounds, then the states will depart within tolerances from the target (and tend asymptotically to it). When the problem concerning an n-dimensional system and an additive cost objective is regular, i.e. when the Hamiltonian of the problem can be uniquely optimized by a control value u^0 depending continuously on the remaining variables (t, x, λ), then a set of $2n$ ordinary differential equations (ODEs) with two-point boundary-value conditions, known as Hamilton (or Hamiltonian) canonical equations (HCEs), has to be solved to obtain the optimal solution. For the linear-quadratic regulator (LQR) with a finite horizon, there exist well-known methods (see for instance Costanza & Neuman, 2009; Sontag, 1998) to transform the boundary-value problem into an initial-value one. In the infinite-horizon, bilinear-quadratic regulator, and change of set-point servo problems, there also exists an attempt to find the missing initial condition for the co-state variable from the data of each particular problem, which allows to integrate the Hamiltonian equations online with the underlying control process (Costanza & Neuman, 2006). For nonlinear systems this line of work is in its beginnings (Costanza & Rivadeneira, 2008, Costanza, Rivadeneira, & Spies, 2009).

Whenever an optimal performance is desired, the bounded-control context may lead to non-regular optimal control problems, for whose solution there are not standard recipes (Athans & Falb, 2006; Qin & Badgwell, 2003; Sontag, 1998; Speyer & Jacobson, 2010). Since the early 1960s, the Pontryagin's maximum principle (PMP) has been at the core of the development of modern optimal control theory (Pontryagin, Boltyanskii, Gamkrelidze, & Mishchenko, 1964) to treat non-regular situations. This paper takes advantage of the relationships between PMP and the Hamilton–Jacobi approaches to the LQR problem. The decisive theoretical finding may be phrased as follows: the optimal solution to a given restricted LQR problem can be generated by saturating the solution of another unrestricted LQR problem, with the same dynamics and cost objective as the original one, but starting at a different initial condition and subject to a quadratic final penalization with a different matrix coefficient. Offline and online schemes were developed to detect this new initial condition and final penalization matrix (Costanza, Rivadeneira, & González, 2014). An online algorithm in this direction is the main contribution of this article from the practical point of view. Since the strategy is intended to work in feedback form when the control is between bounds, then the precise knowledge of the initial condition of the subjacent unrestricted LQR process is not substantial. In fact, in such a context what becomes a priority is the location of

*Corresponding author. Email: tsinoli@santafe-conicet.gov.ar

the time instants where the control saturates/desaturates. Therefore, the proposed scheme updates the final penalization matrix and the saturation/desaturation times referred above, while the total cost is being reduced *via* a straightforward version of the gradient method (Pardalos & Pytlak, 2008). The resulting control will in general be suboptimal with respect to the given initial state, since: (i) the eventual occurrence of state perturbations which deviate the system from the optimal trajectory of the original problem and also that (ii) the number of online updating of the parameters, each consuming some computer time, may not be sufficient to reach their optimal values during the available time horizon. The numerical scheme takes advantage of the online availability of the Riccati matrices that correspond to a range of final penalty parameter values, generated from the solutions to a pair of first-order partial differential equations (PDE) (Costanza & Rivadeneira, 2013). For its simplicity and small computational effort, the online algorithm can be considered as a potential tool to be used in combination with the receding or shrinking horizon policies (Camacho & Bordons, 2004), in an enlarged MPC context that would contemplate strict finite-horizon problems, and may be considered as an alternative to nonlinear programming approaches (Cannon, Liao, & Kouvaritakis, 2008, Rao et al., 2008), which depend on the time and space discretization adopted.

The article has the following structure: after the Introduction, the regular LQR results and the auxiliary matrices that will be used in the sequel are presented. Then the bounded-control version of the problem is described and the main theoretical fact used here is stated. Afterwards the algebraic formulas to be employed in the numerical updating of the parameters are explicitly given. Two applications of the numerical scheme to linearized systems are then illustrated, one of them arising from an industrial problem. The usual Conclusions are given at the end.

2. Equations for regular LQR optimal control problems

The finite-horizon, time-constant formulation of the LQR problem with free final states and unconstrained controls attempts to minimize the (quadratic) cost

$$\mathcal{J}(u) = \int_0^{t_f} [x'(\tau)Qx(\tau) + u'(\tau)Ru(\tau)]\,\mathrm{d}\tau + x'(t_f)Sx(t_f), \tag{1}$$

with respect to all the admissible (here piecewise-continuous) control trajectories $u : [0, t_f] \to \mathbb{R}^m$ of duration t_f, applied to some fixed, finite-dimensional, deterministic plant. Then control strategies affect the $\mathbb{R}^n$-valued states x through some initialized, autonomous, dynamical constraint

$$\dot{x} = Ax + Bu := f(x, u), \quad x(0) = x_0 \neq 0. \tag{2}$$

This will be called a $(A, B, Q, R, S, t_f, \mathbb{R}^m, x_0)$-problem.

The (real, time-constant) matrices in Equations (1) and (2) will be assumed to have the following properties: Q and S are positive-semidefinite $n \times n$ matrices, R is $m \times m$ and positive-definite, A is $n \times n$, B is $n \times m$, and the pair (A, B) is controllable. The expression under the integral is usually known as the 'Lagrangian' L of the cost, namely

$$L(x, u) := x'Qx + u'Ru. \tag{3}$$

Under these conditions the Hamiltonian of the problem, namely the $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ function defined by

$$H(x, \lambda, u) := L(x, u) + \lambda' f(x, u) \tag{4}$$

is known to be regular, i.e. that H is uniquely minimized with respect to u, and this occurs when u takes the explicit control value

$$u^0(x, \lambda) = -\tfrac{1}{2}R^{-1}B'\lambda, \tag{5}$$

(in this case, independently of x), which is usually called 'the H-minimal control'. The 'Hamiltonian' form of the problem (see, for instance, Sontag, 1998) requires then to solve the two-point boundary-value problem for the HCEs

$$\dot{x} = H^0_\lambda(x, \lambda), \quad x(0) = x_0, \tag{6}$$

$$\dot{\lambda} = -H^0_x(x, \lambda), \quad \lambda(t_f) = 2Sx(t_f), \tag{7}$$

where $H^0(x, \lambda)$, usually called the minimized (or control) Hamiltonian, stands for

$$H^0(x, \lambda) := H(x, \lambda, u^0(x, \lambda)), \tag{8}$$

and H^0_λ and H^0_x for the column vectors with i-components $\partial H^0/\partial\lambda_i$, $\partial H^0/\partial x_i$ respectively, i.e. Equations (6) and (7) here take the form

$$\begin{aligned} \dot{x} &= Ax - \tfrac{1}{2}W\lambda, \\ \dot{\lambda} &= -2Qx - A'\lambda, \end{aligned} \tag{9}$$

respectively, with $W := BR^{-1}B'$. It is well known that the solution to the unrestricted regular problem, as posed above, relies in turn on the solution $P(\cdot)$ to the Riccati differential equation (RDE)

$$\dot{P} = PWP - PA - A'P - Q, P(t_f) = S, \tag{10}$$

which establishes a useful relationship between the optimal state $x^*(\cdot)$ and the costate $\lambda^*(\cdot)$ trajectories, namely

$$\lambda^*(t) = 2P(t)x^*(t), \tag{11}$$

and, based on Equation (5), leads to the optimal control trajectory

$$u^*(t) = u^0(x^*(t), \lambda^*(t)) = -R^{-1}B'P(t)x^*(t), \tag{12}$$

or equivalently to the optimal feedback law

$$u_f(t, x) = -R^{-1}B'P(t)x. \tag{13}$$

When the control values are restricted, the global regularity of the Hamiltonian cannot be assured, and therefore the search for the optimal control strategy becomes

more involved, as may be observed in the following sections. Additional relevant objects from the LQR theory will be used in the sequel, for instance, the matrices $\alpha(T,S), \beta(T,S)$, solutions to the following pair of first-order, quasilinear PDE (see Costanza & Neuman, 2009 for details):

$$\alpha_T - \alpha_S M = -\alpha N, \quad \alpha(0,S) = I, \tag{14}$$

$$\beta_T - \beta_S M = -\beta N, \quad \beta(0,S) = 2S, \tag{15}$$

where the subindices denote partial derivatives and the matrix coefficients result

$$M := A'S + SA + Q - SWS, \tag{16}$$

$$N := A - WS. \tag{17}$$

These matrices allow us to calculate, for any unbounded LQR problem, the solution $P(\cdot, t_f, S)$ to its RDE through the formula

$$P(t,t_f,S) = \tfrac{1}{2}\beta(t_f - t, S)[\alpha(t_f - t, S)]^{-1} \quad \forall\, t \in [0, t_f], \tag{18}$$

and in such a case the matrices α and β are also related to the boundary conditions by the following relations:

$$x(0) = \alpha(t_f, S)x(t_f), \lambda(0) = \beta(t_f, S)x(t_f). \tag{19}$$

3. The bounded-control case

The manipulated variable in most of the control systems appearing in practical applications can assume only a bounded set of values. The term 'manipulated' indicates that a person or an instrument assigns a value to a signal generated by physical means, and therefore this value cannot take more than a physically realizable amount. Commonly, the manipulated variable can move inside and on the boundary of some bounded subset of a metric space, then it is natural to assume that the admissible set of control values is a compact subset of $\mathbb{R}$. The qualitative features of optimal control solutions to bounded problems are significantly different from those of unbounded ones (Pontryagin et al., 1964). But questions about how much they actually differ, which classes of problems lead to bang–bang controls, and whether their solutions are just saturations of the optimal trajectories of unbounded problems, are still open. Linearizations are accepted provided fluctuations are small. This principle affects state and control deviations. Particularly, the speed variations cannot be allowed to trespass appropriate bounds, i.e.

$$u(t) \in \mathbb{U} := [u_{\min}, u_{\max}]. \tag{20}$$

This restriction determines the structure of the optimal control problem. When $\mathbb{U}$ is bounded (and closed, as in Equation (20)), then the problem tends to lose regularity; derivatives of the Hamiltonian and the value function are not even guaranteed to exist. The search for solutions to restricted problems most frequently falls in the domains of the Pontryagin's maximum principle (PMP; Pontryagin et al., 1964). However, even when solved, PMP is not flexible enough to treat state perturbations: no optimal feedback laws arise from the application of PMP equations, but only open-loop control strategies.

In this paper the following result (Costanza & Rivadeneira, 2013) will be exploited:

Let us assume that there exists a time instant $t \in (0, t_f)$ where $u^*_{x_0}(t) \in (u_{\min}, u_{\max})$. Then there exists a time interval $I \subset (0, t_f)$ containing t such that the optimal phase trajectory $\{x^*_{x_0}, \lambda^*_{x_0}\}$ of the original $(A, B, Q, R, S, t_f, \mathbb{U}, x_0)$-problem coincides with the optimal phase trajectory $\{\hat{x}, \hat{\lambda}\}$ corresponding to a $(A, B, Q, R, \hat{S}, t_f, \mathbb{R}, \hat{x}_0)$-problem.

In what follows, it will be assumed that there exists just one maximal 'regular' interval $(\tau_1, \tau_2) \subset (0, t_f)$ where the control takes values in $(u_{\min}, u_{\max})$. The generation of a suboptimal control strategy by approximating the unknown parameters $(\hat{S}, \hat{x}_0)$ has already been published in the minimal-control-energy situation (Costanza et al., 2014), and the extension to the general LQR problem was announced in Costanza and Rivadeneira (2013), although there the updating of the unknown parameters still required the simulation of state trajectories. In this paper, explicit algebraic formulas are given for parameters updating, avoiding ODE integrations, and thus reducing the computer time in the process of decreasing the total cost.

3.1. *Algebraic formulas used in the online procedure*

3.1.1. *Auxiliary objects*

The following type of feedback control laws will be frequently used in the sequel leading online to a suboptimal control:

$$\tilde{u}(t) := \begin{cases} u_{\min}, & \forall\, t \in [0, \tau_1), \\ -R^{-1}B'P(t,\tilde{S})x(t), & \forall\, t \in [\tau_1, \tau_2), \\ u_{\max}, & \forall\, t \in [\tau_2, t_f], \end{cases} \tag{21}$$

where $\tilde{u}(t)$ is a short notation for $\tilde{u}_{\tilde{S},\tau_1,\tau_2}(t)$, which will be used when it is necessary to indicate that the feedback law is associated to the parameters $(\tilde{S}, \tau_1, \tau_2)$. The 'seed' strategy will be adopted to start the online iterative method below, and has the same structure, namely

$$u_{\text{seed}}(t) := \begin{cases} u_{\min} & \text{if } -R^{-1}B'P(t,S)x(t) \le u_{\min}, \\ u_{\max} & \text{if } -R^{-1}B'P(t,S)x(t) \ge u_{\max}, \\ -R^{-1}B'P(t,S)x(t) & \text{otherwise.} \end{cases} \tag{22}$$

The state trajectory corresponding to the control u_{seed} and starting at x_0, i.e. $x_{u_{\text{seed}}}$, will be denoted as x_{seed}. Note that through the seed control and state trajectories, simulated offline for the nominal final matrix S, the first values for the saturation times, denoted $\tau_{1,0} \le \tau_{2,0}$, are detected if they

exist. The initial approximation to the unknown penalization matrix $\hat{S}$ will be denoted $\tilde{S}_0 := S$, and consistently for the control, $\tilde{u}_0 := u_{\text{seed}}$. The Hamiltonian matrix of the original problem,

$$\mathbf{H} := \begin{pmatrix} A & -\frac{W}{2} \\ -2Q & -A' \end{pmatrix}, \tag{23}$$

and the associated fundamental matrix

$$\mathbf{V}(t) := \mathrm{e}^{\mathbf{H}t}, \tag{24}$$

will also be employed in devising algebraic formula for the partial derivatives of the cost. The matrix $\mathbf{V}(t)$ is $2n \times 2n$. For convenience, its $n \times n$ partition is denoted as

$$\begin{pmatrix} \mathbf{V}_1(t) & \mathbf{V}_2(t) \\ \mathbf{V}_3(t) & \mathbf{V}_4(t) \end{pmatrix} := \mathbf{V}(t), \tag{25}$$

which can also be expressed in terms of the auxiliary matrices α, β, and their derivatives since, Costanza and Neuman (2009)

$$\mathbf{V} = (\mathrm{e}^{\mathbf{H}T})^{-1} = \begin{pmatrix} \alpha - \alpha_S S & \alpha_S/2 \\ \beta - \beta_S S & \beta_S/2 \end{pmatrix}^{-1}. \tag{26}$$

In what follows $P(t, \tilde{S})$ will denote the solution to the RDE (10) for π, with final condition $\pi(t_\mathrm{f}) = \tilde{S}$. When the value of $\tilde{S}$ is clear in the text, the notation may simplify from $P(t, \tilde{S})$ to $P(t)$. The following identity will also be used:

$$\begin{aligned} \frac{\partial P(t,\tilde{S})}{\partial \tilde{S}} &= \frac{\partial [(1/2)\beta(t_\mathrm{f}-t,\tilde{S})[\alpha(t_\mathrm{f}-t,\tilde{S})]^{-1}]}{\partial \tilde{S}} \\ &= \tfrac{1}{2}[\beta_S \alpha^{-1} - \beta\alpha^{-1}\alpha_S\alpha^{-1}](t_\mathrm{f}-t,\tilde{S}) \\ &= \tfrac{1}{2}[\beta_S - 2P(t,\tilde{S})\alpha_S]\alpha^{-1}. \end{aligned} \tag{27}$$

The 'saturated' fundamental matrix:

$$\Psi(t,\tau) := \int_\tau^t \mathrm{e}^{A(t-\sigma)}\,\mathrm{d}\sigma = \mathrm{e}^{At}\int_\tau^t \mathrm{e}^{-A\sigma}\,\mathrm{d}\sigma, \tag{28}$$

and the related matrices

$$\check{\Psi}(t,\tau) := \int_\tau^t \mathrm{e}^{A'(\sigma-\tau)} Q\, \mathrm{e}^{A(\sigma-\tau)}\,\mathrm{d}\sigma, \tag{29}$$

$$\hat{\Psi}(t,\tau) := \int_\tau^t \Psi'(\sigma,\tau) Q\, \mathrm{e}^{A(\sigma-\tau)}\,\mathrm{d}\sigma \tag{30}$$

will also be needed in the sequel. These matrices are calculated and interpolated offline. So it is assumed that, in real-time applications, they will be available as functions of their two variables (t, τ), in the range $[0, t_\mathrm{f}] \times [0, t_\mathrm{f}]$.

Since $u(t) \equiv u_{\min}$ in $[0, \tau_1]$, then the state at time τ_1 results

$$x(\tau_1) = \mathrm{e}^{At}x_0 + \Psi(\tau_1, 0)Bu_{\min}. \tag{31}$$

Here it is important to note that the state at τ_2 may be calculated in two equivalent ways: (i) either from the Hamiltonian flow, since for each $t \in [\tau_1, \tau_2)$ the control is $-R^{-1}B'P(t,\tilde{S})x(t)$ and the costate (corresponding to this piece of a regular trajectory), denoted $\tilde{\lambda}(t)$, is $\tilde{\lambda}(t) = 2P(t,\tilde{S})x(t)$, and therefore

$$\begin{pmatrix} x(t) \\ \tilde{\lambda}(t) \end{pmatrix} = \mathbf{V}(t-\tau_1)\begin{pmatrix} x(\tau_1) \\ \tilde{\lambda}(\tau_1) \end{pmatrix}, \tag{32}$$

implying that

$$\begin{aligned} x(\tau_2) &= \mathbf{V}_1(\tau_2-\tau_1)x(\tau_1) + \mathbf{V}_2(\tau_2-\tau_1)\tilde{\lambda}(\tau_1) \\ &= (\mathbf{V}_1(\tau_2-\tau_1) + 2\mathbf{V}_2(\tau_2-\tau_1)P(\tau_1,\tilde{S}))x(\tau_1), \end{aligned} \tag{33}$$

or (ii) $x(t)$ should also coincide with the state of some process having the same final penalization $\tilde{S}$ and starting at some initial state $\tilde{x}(0)$, i.e.

$$\begin{pmatrix} x(t) \\ \tilde{\lambda}(t) \end{pmatrix} = \mathbf{V}(t)\begin{pmatrix} \tilde{x}(0) \\ 2P(0,\tilde{S})\tilde{x}(0) \end{pmatrix}, \tag{34}$$

$$x(\tau_2) = (\mathbf{V}_1(\tau_2) + 2\mathbf{V}_2(\tau_2)P(0,\tilde{S}))\tilde{x}(0). \tag{35}$$

In what follows, the arguments t of $\mathbf{V}(t)$ will also be omitted when the context is clear. The state at the final time t_f is

$$x(t_\mathrm{f}) = \mathrm{e}^{A(t_\mathrm{f}-\tau_2)}x(\tau_2) + \Psi(t_\mathrm{f},\tau_2)Bu_{\max}. \tag{36}$$

3.1.2. *The partial derivatives of the cost*

It is known (Dhamo & Tröltzsch, 2011) that the total cost $\mathcal{J}(\tilde{u})$ is differentiable as a function of the variables $(\tilde{S}, \tau_1, \tau_2)$. The total cost is time-partitioned here for convenience:

$$J(\tilde{S},\tau_1,\tau_2) := \mathcal{J}(\tilde{u}) := J_1 + J_2 + J_3 + J_4, \tag{37}$$

$$J_1 := \int_0^{\tau_1} L(x_{\tilde{u}}(t),\tilde{u}(t))\,\mathrm{d}t = Ru_{\min}^2\tau_1 + \int_0^{\tau_1} x'(t)Qx(t)\,\mathrm{d}t, \tag{38}$$

$$\begin{aligned} J_2 := \int_{\tau_1}^{\tau_2} L\,\mathrm{d}t &= x'(\tau_1)P(\tau_1,\tilde{S})x(\tau_1) \\ &- x'(\tau_2)P(\tau_2,\tilde{S})x(\tau_2), \end{aligned} \tag{39}$$

$$J_3 := \int_{\tau_2}^{t_\mathrm{f}} L\,\mathrm{d}t = Ru_{\max}^2(t_\mathrm{f}-\tau_2) + \int_{\tau_2}^{t_\mathrm{f}} x'(t)Qx(t)\,\mathrm{d}t, \tag{40}$$

$$J_4 := x'(t_\mathrm{f})Sx(t_\mathrm{f}). \tag{41}$$

Based on the preliminary formulas, the partial derivatives of each partial cost can be expressed as

$$D_{\tau_1}J_1 = Ru_{\min}^2 + x'(\tau_1)Qx(\tau_1), \tag{42}$$

where $x(\tau_1)$ and $x'(\tau_1)$ should be replaced by their corresponding expressions from Equation (31), as in all

succeeding partial derivatives. Next,

$$D_{\tau_1}J_2 = 2x'(\tau_1)P(\tau_1,\tilde{S})\dot{x}(\tau_1) + x'(\tau_1)\dot{P}(\tau_1,\tilde{S})x(\tau_1) - \cdots - \frac{\partial}{\partial\tau_1}(x'(\tau_2)P(\tau_2,\tilde{S})x(\tau_2)) \quad (43)$$

$$= 2x'(\tau_1)P(\tau_1)[Ax(\tau_1) + Bu_{\min}] + \cdots + x'(\tau_1)[P(\tau_1)WP(\tau_1) - P(\tau_1)A - A'P(\tau_1) - Q]x(\tau_1), \quad (44)$$

since $(\partial/\partial\tau_1)[x'(\tau_2)P(\tau_2)x(\tau_2)] = 2x'(\tau_2)P(\tau_2)(\partial x(\tau_2)/\partial\tau_1)$, and, from Equation (35), $\partial x(\tau_2)/\partial\tau_1 = 0$.

$$D_{\tau_1}J_3 = D_{\tau_1}\int_{\tau_2}^{t_f} x'(t)Qx(t)\,dt = \int_{\tau_2}^{t_f} 2x'(t)Q\frac{\partial x(t)}{\partial\tau_1}\,dt = 0, \quad (45)$$

since inside the integral of the state $x(t) = e^{A(t-\tau_2)}x(\tau_2) + \Psi(t_f,\tau_2)Bu_{\max}$, $\partial x(t)/\partial\tau_1 = e^{A(t-\tau_2)}\ \partial x(\tau_2)/\partial\tau_1$, and $\partial x(\tau_2)/\partial\tau_1 = 0$.

$$D_{\tau_1}J_4 = 2x'(t_f)S\frac{\partial x(t_f)}{\partial\tau_1} = 2x'(t_f)S\,e^{A(t_f-\tau_2)}\frac{\partial x(\tau_2)}{\partial\tau_1} = 0. \quad (46)$$

Similarly, the derivatives with respect to τ_2 are

$$D_{\tau_2}J_1 = 0, \quad (47)$$

$$D_{\tau_2}J_2 = -[2x'(\tau_2)P(\tau_2,\tilde{S})\dot{x}(\tau_2) + x'(\tau_2)\dot{P}(\tau_2)x(\tau_2)] = -\{2x'(\tau_2)P(\tau_2)(Ax(\tau_2) + Bu_{\max}) + \cdots + x'(\tau_2)P(\tau_2)WP(\tau_2) - P(\tau_2)A - A'P(\tau_2) - Q)x(\tau_2)\}, \quad (48)$$

where Equations (33) and (31) should be used to replace $x(\tau_2)$, as in the next partial derivative:

$$D_{\tau_2}J_3 = -[Ru_{\max}^2 + x'(\tau_2)Qx(\tau_2)], \quad (49)$$

$$D_{\tau_2}J_4 = 2x'(t_f)S\frac{\partial x(t_f)}{\partial\tau_2} = -2x'(t_f)S\,e^{A(t_f-\tau_2)}[WP(\tau_2,\tilde{S})x(\tau_2) + Bu_{\max}], \quad (50)$$

since, from Equation (36) $\partial x(t_f)/\partial\tau_2 = -A\,e^{A(t_f-\tau_2)}x(\tau_2) + e^{A(t_f-\tau_2)}\dot{x}(\tau_2) - e^{A(t_f-\tau_2)}Bu_{\max}$, and $\dot{x}(\tau_2) = (A - WP(\tau_2,\tilde{S}))x(\tau_2)$. The derivatives with respect to the elements $\tilde{S}_{ij}$ of $\tilde{S}$, globally denoted as $D_{\tilde{S}}$, are

$$D_{\tilde{S}}J_1 = 0, \quad (51)$$

$$D_{\tilde{S}}J_2 = x'(\tau_1)\left[\frac{\partial P(\tau_1)}{\partial\tilde{S}} - 4(\mathbf{V}_1 + 2\mathbf{V}_2P(\tau_1))'P(\tau_2)\mathbf{V}_2\frac{\partial P(\tau_1)}{\partial\tilde{S}} - \cdots - (\mathbf{V}_1 + 2\mathbf{V}_2P(\tau_1))'\frac{\partial P(\tau_2)}{\partial\tilde{S}}(\mathbf{V}_1 + 2\mathbf{V}_2P(\tau_1))\right]x(\tau_1), \quad (52)$$

$$D_{\tilde{S}}J_3 = \int_{\tau_2}^{t_f} 2x'(t)Q\frac{\partial x(t)}{\partial\tilde{S}}\,dt = 4[x'(\tau_2)\check{\Psi}(t_f,\tau_2) + u_{\max}B'\hat{\Psi}(t_f,\tau_2)]\mathbf{V}_2(\tau_2-\tau_1)\frac{\partial P(\tau_1,\tilde{S})}{\partial\tilde{S}}x(\tau_1), \quad (53)$$

after replacing $x'(t) = x'(\tau_2)\,e^{A'(t-\tau_2)} + B'\Psi'(t,\tau_2)u_{\max}$, and expanding $\partial x(t)/\partial\tilde{S} = e^{A(t-\tau_2)}(\partial x(\tau_2)/\partial\tilde{S}) = e^{A(t-\tau_2)}\ 2\mathbf{V}_2(\tau_2-\tau_1)(\partial P(\tau_1,\tilde{S})/\partial\tilde{S})x(\tau_1)$. Finally, after similar manipulations,

$$D_{\tilde{S}}J_4 = 2x'(t_f)S\frac{\partial x(t_f)}{\partial\tilde{S}} = 4x'(t_f)S\,e^{A(t_f-\tau_2)}\mathbf{V}_2(\tau_2-\tau_1)\frac{\partial P(\tau_1,\tilde{S})}{\partial\tilde{S}}x(\tau_1). \quad (54)$$

3.1.3. *Updating the parameters*

First approximations $\tau_{1,0}$ and $\tau_{2,0}$ to the optimal saturation points τ_1 and τ_2 become available after (offline) simulating the state trajectory x_{seed}. A subdivision of the time-horizon in 'sampling times' of the form $t_0 = 0 < t_1 < t_2 < \cdots < t_N = T$ is adopted to make possible intermediate calculations, updating parameters, and deciding changes in the control strategy. For $t \in [0,t_1]$ the control is set to

$$u(t) \equiv u_{\text{seed}}(t). \quad (55)$$

Then, during this initial sampling interval (associated with $k = 0$) and through, the parameters $(\tilde{S},\tau_1,\tau_2)$ are updated to construct successive control strategies $\tilde{u}_{k,j}$, $j = 1,2,\ldots$ that decrease the value of the total cost:

$$\mathcal{J}(\tilde{u}_{k,j+1}) \le \mathcal{J}(\tilde{u}_{k,j}) \le \cdots \mathcal{J}(u_{\text{seed}}), \quad j = 1,2,\ldots \quad (56)$$

according to the prescriptions of the simplest gradient method:

$$\tilde{S}_{k,j} := \tilde{S}_{k,j-1} - \gamma_S\frac{\partial J}{\partial\tilde{S}}(\tilde{S}_{k,j-1},\tau_{1,k},\tau_{2,k}), \quad j = 1,2,\ldots. \quad (57)$$

The last updating of $\tilde{S}_k$ that can be computed during the sampling interval is denoted

$$\tilde{S}_{k+1} \approx \lim_j \tilde{S}_{k,j}, \quad (58)$$

and on the same lines,

$$\tau_{i,k+1} \approx \lim_j\left(\tau_{i,k} - \gamma_{\tau_i}\frac{\partial J}{\partial\tau_i}\right)_j, \quad i = 1,2, \quad (59)$$

where $\gamma_S,\gamma_{\tau_1},\gamma_{\tau_2}$ are appropriate constants, tuned by the user for each experiment. During the next sampling interval $(t_{k+1},t_{k+2}]$, the control is set to

$$u(t) \equiv \tilde{u}_{\tilde{S}_{k+1},\tau_{1,k+1},\tau_{2,k+1}}(t). \quad (60)$$

4. Applications and numerical results

4.1. A one-dimensional example

The first case study of this paper is the optimal control problem defined by the following objects:

$$\begin{aligned} &\dot{x}(t) = u(t), \quad 0 \le t \le 1, \\ &x(0) = 1 - e, \\ &u(t) \in [1.44, 2] \subset \mathbb{R}, \\ &J(u) = \int_0^1 (x^2 + u^2)\mathrm{d}t + 13\,[x(1)]^2\,. \end{aligned} \tag{61}$$

The optimal solution is given in Costanza and Rivadeneira (2013). A slightly different version of this problem, aimed to minimize only the control energy, was presented in Troutman (1996). For any real constant C, this system can also be considered as a linearized version of the dynamics

$$\dot{z}(t) = U(t) + C, \tag{62}$$

with state $z(t)$, control $U(t)$, steady-state $z_{SS}(t) \equiv 0$, equilibrium control $U_{SS}(t) = -C$, and deviations $x(t) := z(t) - z_{SS}(t), u(t) := U(t) - U_{SS}(t)$.

The gradient method was first tried offline (without discretizing the time-horizon into sampling periods), to obtain: $\tilde{S} = 9.60305$, $\tau_1 = 0.09289$, $\tau_2 = 0.62095$, $J_{off} = 3.73091 \approx J^* = 3.7309$ (Costanza & Rivadeneira, 2013). It was also implemented online, as described above, by adopting a fixed sampling period $\Delta t_k = t_{k+1} - t_k = 0.025$ and allowing for a maximum of 30 iterations inside each Δt_k. The total cost obtained was $J_{on} = 3.73123$, slightly higher than $J_{off} \approx J^*$. The resulting online control trajectory, and the evolution of the required parameters $\tau_1, \tau_2, \tilde{S}$, updated at each sampling-time, are depicted in Figures 1 and 2, respectively.

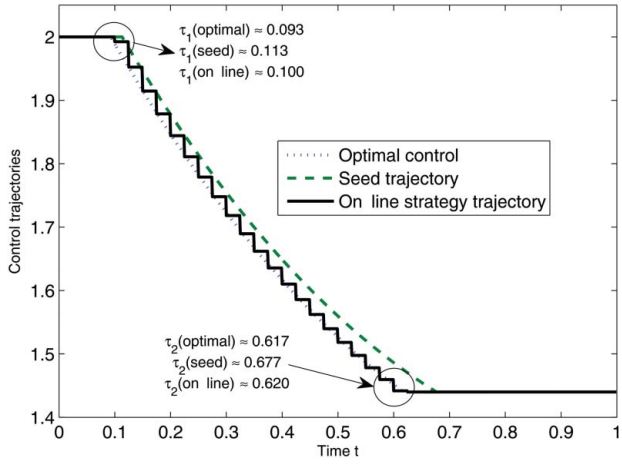


Figure 1. Control strategy resulting from the online application of the gradient method to the one-dimensional example.

4.2. A typical linearized model situation: the rolling mill

4.2.1. The nonlinear first-order PDE setup

The second case study models a rolling mill described in Hearns and Grimble (2010), whose (infinite dimensional) dynamics (from a standard energy balance) obeys the following first-order PDE

$$\frac{\partial \theta}{\partial t} = -V\frac{\partial \theta}{\partial z} + a(\theta_a - \theta) + b(\theta_a^4 - \theta^4), \tag{63}$$

where $\theta(t, z)$ is the temperature of the metallic strip at time t and location z in the trend, $V(t)$ is the linear speed of the strip, and θ_a is the ambient temperature (assumed constant in this set up). The coefficients a, b weigh the rate of heating due to conduction and radiation, respectively. The system

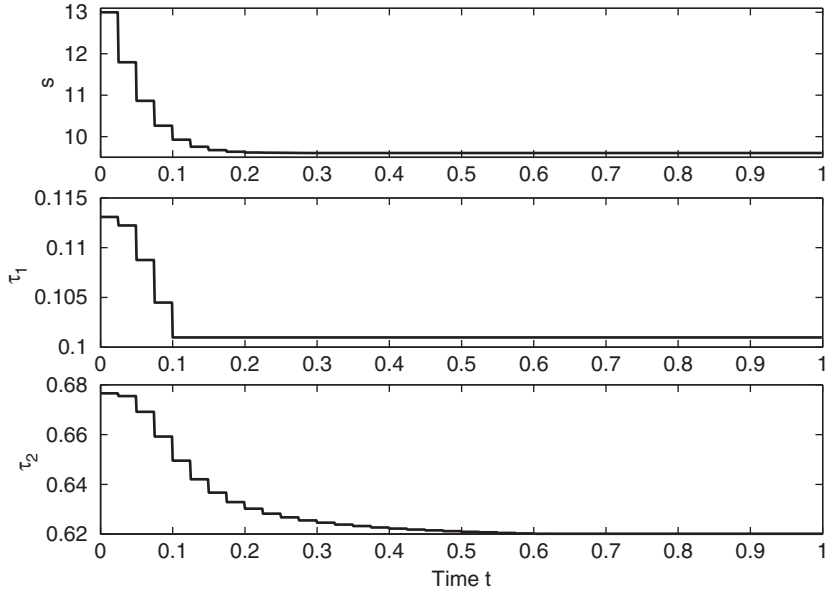


Figure 2. Evolution of the parameters $\tilde{S}, \tau_1, \tau_2$, during the online application of the gradient method.

is simplified by neglecting radiation (small b), and by supposing that the temperature will stay around the equilibrium profile

$$\theta_{SS}(z) = \theta_a + (\theta_0 - \theta_a)\exp\left(-\frac{az}{V_0}\right), \tag{64}$$

which is the solution to Equation (63) with $b = 0$, $\partial\theta/\partial t = 0$, $V(t) \equiv V_0$, and $\theta_0 := \theta_{SS}(0)$, some appropriate constant characterizing each physical set up. The following definitions

$$\Delta\theta(t,z) := \theta(t,z) - \theta_{SS}(z), \quad u(t) := V(t) - V_0 \tag{65}$$

allow one to approximately express the dynamics of the fluctuations through the 'linearized' version of Equation (63), namely

$$\frac{\partial\Delta\theta}{\partial t} = -V_0\frac{\partial\Delta\theta}{\partial z} - a\Delta\theta + \left[\frac{a}{V_0}(\theta_0 - \theta_a)\exp\left(-\frac{az}{V_0}\right)\right]u, \tag{66}$$

after neglecting the term $u\partial\Delta\theta/\partial z$, on the argument that it is the product of two 'small' quantities.

Remark 1 The linear approximation of the original problem implicitly supposes not only that deviations $\Delta\theta(t,z)$ from the steady-state temperature $\theta_{SS}(z)$ are small, but also that the manipulated variable $V(t)$ will also be near the steady-state velocity V_0. These assumptions imply, especially, that the control variable $u(t)$ should not be allowed to take unbounded values, even if their physical realization are possible, because the linearized model will risk to depart too much from the original dynamics.

4.2.2. *The z-discretization approach leading to a finite-dimensional linear control system*

From the control theory perspective, the state in Equation (66) is at each time t the z-function $\Delta\theta(t,\cdot)$. This, in principle, makes the system under study infinite-dimensional, whose treatment is out of the scope of this paper. An n-dimensional approximation has then been constructed by discretizing the z-variable in the form:

$$z_i := (i-1)h, \quad i = 1,\ldots,n, \tag{67}$$

next by defining n state variables x_i (or equivalently a vector state variable $x(\cdot)$ with values $x(t)$ in $\mathbb{R}^n$),

$$x_i(t) := \Delta\theta(t,z_i), \quad i = 1,\ldots,n, \tag{68}$$

$$x(t) := (x_1(t), x_2(t), \ldots, x_n(t))', \tag{69}$$

and finally by approximating the z-partial derivative by some appropriate linear combination of the function $\Delta\theta(t,\cdot)$ evaluated at the discretized values z_i, for instance,

$$\frac{\partial\Delta\theta}{\partial z}(t,z_i) \approx \frac{x_{i+1}(t) - x_i(t)}{h}, \quad i = 1,\ldots,n-1, \tag{70}$$

$$\frac{\partial\Delta\theta}{\partial z}(t,z_n) \approx \frac{x_n(t) - x_{n-1}(t)}{h}. \tag{71}$$

After such manipulations the following structure of a linear control system is obtained

$$\dot{x} = Ax + Bu, \tag{72}$$

where the $n \times n$ matrix A and the column n-vector B take the form:

$$A = (a_{ij}) : \begin{cases} a_{ii} = \dfrac{V_0}{h} - a, & a_{i,i+1} = -\dfrac{V_0}{h}, \\ & i = 1,\ldots,n-1, \\ a_{n,n-1} = \dfrac{V_0}{h}, & a_{nn} = -\left(a + \dfrac{V_0}{h}\right), \\ \text{all remaining} \\ \text{elements equal to 0,} \end{cases} \tag{73}$$

$$B = (b_i) = \frac{a}{V_0}(\theta_0 - \theta_a)\exp\left(-\frac{az_i}{V_0}\right), \quad i = 1,\ldots,n. \tag{74}$$

The eigenvalues of the matrix A are dominated by the relation between the heat gained at each position by convection versus the heat extracted at that point by the environment, implicit in the term $V_0/h - a$ which appears in the main diagonal, except in its last element. Now, from one side, the free evolution has to be stable to keep any physical meaning in the equations (the temperature cannot grow forever). But, if control has to be relevant to increase stability, it is appropriate to explore those situations near where the system might lose stability (for instance, due to environmental perturbations). With this contradictory objectives in mind, the following values for the parameters were investigated

$$V_0 = h = 1, \quad a = 1.001. \tag{75}$$

The discretized, ODE version (72) of Equation (66) was numerically confirmed to be an acceptable approximation.

4.2.3. *Numerical simulation of the online strategy*

The initial state $x_0 = x(0)$ used for simulation of the system defined by Equations (72)–(74) was

$$x_i(0) = 100\sin\left(\frac{2\pi z_i}{10}\right), \quad i = 1,\ldots,n, \tag{76}$$

with the following values for the reference temperatures (in °C)

$$\theta_a = 20, \quad \theta_0 = 700. \tag{77}$$

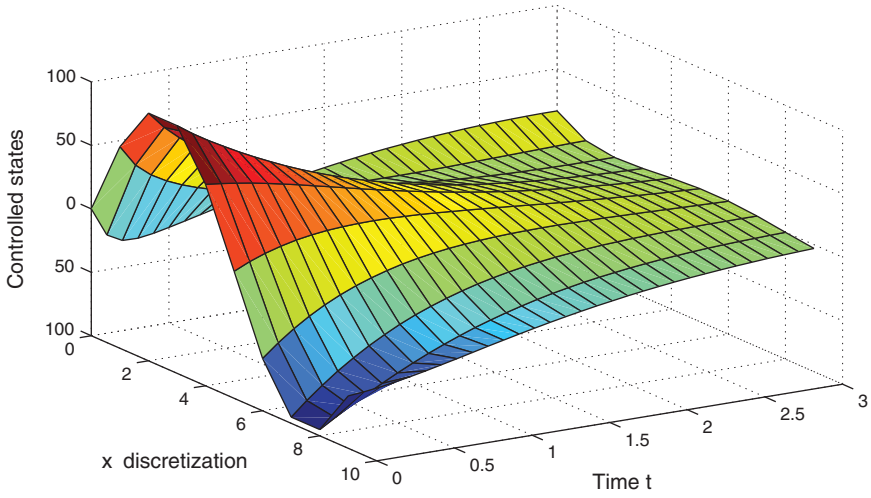


Figure 3. Evolution of the states under the online control strategy, after a sinusoidal initial profile.

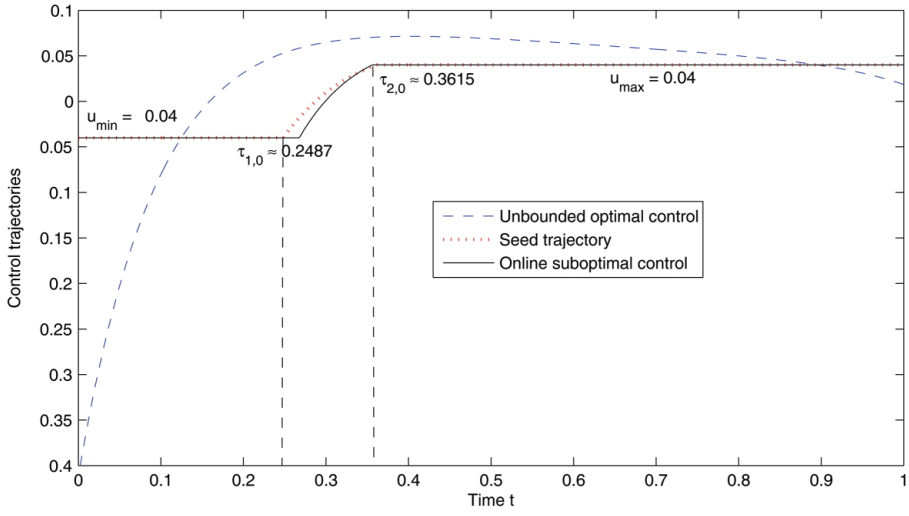


Figure 4. Unbounded optimal control, restricted seed control, and the resulting suboptimal control after applying the gradient method.

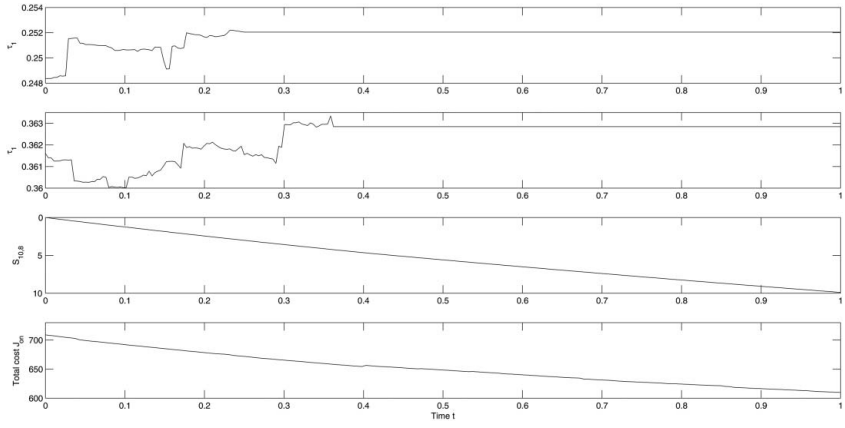


Figure 5. Evolution of some updated parameters and their corresponding total costs.

The cost objective of the LQR problem was given the following parameters:

$$t_f = 1, \quad Q = 0.1I_{10}, \quad R = 500, \quad S = 50I_{10}, \tag{78}$$

and the bounds imposed on control values were

$$[u_{\min}, u_{\max}] = [-0.04, 0.04]. \tag{79}$$

After simulating the seed control and state trajectories, it was found that

$$\tau_{1,0} = 0.2487, \quad \tau_{2,0} = 0.3615, \quad J_{\text{seed}} = 721. \tag{80}$$

Results are shown in Figures 3–5. In Figure 3 the evolution of the states under control is shown. All the components of the state tend to equilibrium, as expected. The unbounded optimal control is shown in Figure 4, together with the seed control trajectory and the suboptimal bounded control resulting from applying the online strategy with fixed sampling intervals $\Delta t = 0.1$. The updating of the parameters τ_1 and τ_2, together with $\tilde{S}_{10,8} = \tilde{S}_{8,10}$ (just one non-diagonal coefficient of $\tilde{S}$ for illustration), and the evolution of the total cost associated with those parameters, are depicted in Figure 5. The relevant final values for the parameters were

$$\begin{aligned} \tau_1 = 0.2590, \quad \tau_2 = 0.3623, \quad J_{\text{on}} = 601, \\ \text{diag}(\tilde{S}) = (50.0, \ldots, 49.9944, 53.5244, 68.3213), \\ \text{subdiag}(\tilde{S}) = (0, \ldots, -0.7454, 0.1887), \\ \text{subsubdiag}(\tilde{S}) = (0, \ldots, -10.5406). \end{aligned} \tag{81}$$

The reduction in the total cost J_{on} with respect to the cost of the seed strategy J_{seed} was 16.78%, mostly due to the high values assigned to R and S in the original cost objective formulation.

5. Conclusions

An efficient online scheme to calculate suboptimal control strategies for linearized models has been presented. Linearized models are thought as approximate versions of real-life nonlinear systems. Linearizations are accurate provided deviations from equilibrium values are small. This context imposes restrictions on the control values, and then normally the Pontryagin approach is required to solve the bounded LQR problem, instead of the more agreeable Hamilton–Jacobi theory. Although feedback laws may be preferred in practice, when perturbations are expected to appear, a closed-loop control is in general suboptimal when there exist constraints in the manipulated variable. With these limitations in mind, an efficient online algorithm is devised to approximate the open-loop optimal control *via* feedback, based on recent theoretical results, and its features are illustrated when applied to two case-studies. The resulting strategies are quite different from the saturated form of the optimal control corresponding to the unrestricted problem with same parameters and initial condition, which is used here just as a first approximation, and so labelled as a 'seed' strategy. Such a seed feedback is often naively adopted in Engineering practice during the whole optimization period, although it has been shown that it is far from optimal. Reductions in total cost, in both examples, validate this assertion. It should be acknowledged that the alternative procedure proposed here will also be suboptimal. This is because the application of PMP to obtain the optimal solution is essentially an offline calculation leading to an open-loop recipe, and if any deviation from the optimal solution occurs (by mistake or by ignorance), then optimality will immediately be lost, no matter the subsequent effort. However, when the PMP solution was not previously found, or when only the 'seed' strategy is available, or when state perturbations appear in a real process-control situation; then no more than a suboptimal performance better than the seed's one can be expected. The online updating of the parameter $\tilde{S}$ and the saturation times τ_i, as long as the total cost is reduced (guaranteed by the gradient method), will clearly improve the seed strategy as time evolves. This new scheme will result in the optimal strategy only when: (i) the right (optimal) $\tilde{S}$ value is reached before the Riccati gain $P(\tilde{S})$ has to be applied, and (ii) no state perturbations occur. As a consequence, the stability of the method is guaranteed since the cost is not allowed to increase, and it is bounded from below. Some positive features of the new online proposed strategy are as follows:

- The method is based on theoretical results ensuring that the hidden final penalization $\hat{S}$ and the appropriate (two at the most) saturation times τ_1 and τ_2 are the critical objects to be ascertained.
- It takes advantage of the availability of α and β as functions of $(T - t, S)$, and consequently on the possibility of generating Riccati matrices $P(t, T, \tilde{S})$ online by simple algebraic manipulations, as $\tilde{S}$ is updated; i.e. the RDE does not need to be solved for any value of $\tilde{S}$, not even offline.
- The control in Equation (21) is given in feedback form, and therefore the algorithm is unaffected by state perturbations due to fluctuations in environmental conditions.
- The updating of parameters $(\tilde{S}, \tau_1, \tau_2)$ is performed via the gradient of the cost of the process, and this cost is calculated by simple algebraic formulas instead of by predicting state, control, and cost trajectories by ODE integrations, as in most 'predictive control' techniques. This reduces the computational effort and allows for updating in shorter sampling intervals.
- Another conceptual difference with currently available approaches is that there exists a unique matrix $\hat{S}$ to look for in treating each LQR problem. This allows for further reduction on the computing effort,

since there is no need for updating Riccati equations through receding horizon schemes.

- It is under exploration the online generation of the matrices α and β involved in the calculation of the optimal feedback gain at each sampling time. This step will improve the applicability of the algorithm to large-dimensional processes, especially to those governed by PDE.

Acknowledgements

We thank Prof. Michael Grimble, at the Industrial Control Centre, Strahclyde University, for introducing us to the problem of temperature control in a metallic strip leaving a rolling mill.

References

Athans, M., & Falb, P.L. (2006). *Optimal control: An introduction to the theory and its applications*. New York: Dover.

Camacho, E. F., & Bordons, C. (2004). *Model predictive control* (2nd ed.). London: Springer.

Cannon, M., Liao, W., & Kouvaritakis, B. (2008). Efficient MPC optimization using Pontryagin's minimum principle. *International Journal of Robust and Nonlinear Control*, *18*, 831–844.

Costanza, V., & Neuman C. E. (2006). Optimal control of nonlinear chemical reactors via an initial-value Hamiltonian problem. *Optimal Control Applications & Methods*, *27*, 41–60.

Costanza, V., & Neuman, C. E. (2009). Partial differential equations for missing boundary conditions in the linear-quadratic optimal control problem. *Latin American Applied Research*, *39*, 207–212.

Costanza, V., & Rivadeneira, P. S. (2008). Finite-horizon dynamic optimization of nonlinear systems in real time. *Automatica*, *44*, 2427–2434.

Costanza, V., & Rivadeneira, P. S. (2013). Optimal saturated feedback laws for LQR problems with bounded controls. *Computational and Applied Mathematics*, *32*, 355–371.

Costanza, V., Rivadeneira, P. S., & González, A.H. (2014). Minimizing control-energy in a class of bounded-control LQR problems. *Optimal Control Applications & Methods*, *35*, 361–382.

Costanza, V., Rivadeneira, P. S., & Spies, R. D. (2009). Equations for the missing boundary values in the Hamiltonian formulation of optimal control problems. *Journal of Optimization Theory and Applications*, *149*, 26–46.

Dhamo, V., & Tröltzsch, F. (2011). Some aspects of reachability for parabolic boundary control problems with control constraints. *Computational Optimization and Applications*, *50*, 75–110.

Hearns, G., & Grimble, M. J. (2010, June 30–July 02). *Temperature control in transport delay systems*. Paper presented at the 2010 American control conference, Baltimore, MD, USA.

Pardalos P., & Pytlak, R. (2008). *Conjugate gradient algorithms in nonconvex optimization*. New York: Springer.

Pontryagin, L.S., Boltyanskii, V. G., Gamkrelidze, R. V., & Mishchenko, E. F. (1964). *The mathematical theory of optimal processes*. New York: Macmillan.

Qin, S. J., & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, *11*, 733–764.

Rao, A. V., Benson, D. A., Huntington, G. T., Francolin, C., Darby, C. L., & Patterson, M. A., *User's manual for GPOPS: A MATLAB package for dynamic optimization using the gauss pseudospectral method*, University of Florida Report, August 2008.

Sontag, E. D. (1998). *Mathematical control theory* (2nd ed.). New York: Springer.

Speyer, J. L., & Jacobson, D. H. (2010). *Primer on optimal control theory*. Philadelphia, PA: SIAM Books.

Troutman, J. L. (1996). *Variational calculus and optimal control*. New York: Springer.